Report on the EOS Aura Algorithm Working Group Meeting, June 26–26, Pasadena, CA

Nathaniel Livesey

July 14, 2000

1 Introduction

This document is a report of the EOS Aura Algorithm Working Group meeting which took place at the Raytheon SDSIO facility adjacent to JPL in Pasadena California. The main aim of the meeting was to plan in detail the first algorithm intercomparison exercise for the EOS Aura instruments. In addition to this planning, general discussion of algorithm issues took place, along with 15 minute AGU style presentations, and algorithm and/or status overviews from each instrument team.

2 Algorithm status and contributed talks

After general introductory discussion, each of the four instrument teams (HIRDLS, MLS, OMI and TES) gave a half hour presentation covering issues such as instrument, algorithm theoretical basis, algorithm and software status.

2.1 Contributed talks

This was followed by a series of 15 minute contributed talks.

Hugh Pumphrey Averaging in information space: a recipe for dealing with weak signals in the presence of strong ones.

This talk described the algorithm the MLS team plans to use for species who's signal to noise ratio is such that Level 2 data will not be useful for scientific study. For these species, Level 3 quantities such as monthly maps or zonal means will be the most useful product. The talk introduced the algorithm, and illustrated it with a simplified simulation of the MLS measurement system.

Robert Voors An ozone profile retrieval algorithm for OMI.

This talk gave an overview of the algorithm the OMI team intend to implement for the ozone profile retrieval (as distinct from column abundances) for this nadir-viewing UV/VIS spectrometer. Emphasis was put on the daily global coverage together with a small (13$\times$24km) ground pixel size, which makes monitoring of the ozone layer possible .

William Read Algorithm approach for UARS MLS upper tropospheric humidity.

This talk reviewed the algorithm used to retrieve upper tropospheric water vapor from UARS MLS data. This measurement is very non-linear, and emphasis was made of the care with which the algorithm used was devised and tested.

Arun Gopalan Validation plans for OMI using data assimilation.

This brief talk outlined future plans for collaboration between the US OMI science team and the Goddard Data Assimilation Office.

Tilman Steck Profile retrieval in the ENVISAT MIPAS Level 2 off line processor.

MIPAS is a limb sounding Fourier transform spectrometer, similar to TES in many respects, MIPAS however focuses on the stratosphere, and only has a limb viewing mode.

Tilman presented a way of applying Tikhonov regularization using the *a priori* covariance matrix to approximate the discrete first derivative operator. This allows a smoothing constraint without a strong constraint to the *a priori* profile. He used diagnostics of *a posteriori* covariance matrices, averaging kernels and degrees of freedom to assess the effectiveness of the constraints.

Clive Rodgers Intercomparison of remote sounders.

This talk considered ways in which data from remote sounders could be compared, in addition to the traditional 'profile' comparison. One possible technique is to identify combinations of radiance measurements in each instrument which convey information on the same region of the atmosphere (i.e. have 'similar' weighting functions).

Clive was unable to attend, this talk was given by Kevin Bowman.

2.2 Introduction to the intercomparison exercise

The contributed talks were followed by an overview from each instrument team on what plans they have for testing with simulated data.

HIRDLS plans to use data from the MOZART model. These are the same data planned for the algorithm intercomparison exercise (discussed below.)

OMI will be using ozone assimilation for testing and simulation, and may use MOZART for NO_2 etc.

MLS will be using data from the SLIMCAT model provided by the University of Edinburgh. One month's worth of data will be generated for use in testing of the algorithms (notably Level 3 which requires a long input dataset.)

TES will also be using the MOZART dataset.

3 Planning for the Algorithm intercomparison exercise

3.1 Fundamental issues

3.1.1 Aims of the exercise

Discussion started by considering what we as a group hope to yield from such an exercise, over and above the work we would be doing independently testing our own algorithms.

- To increase our knowledge of the expected performance of our algorithms and the accuracy and precision of our data.
- To get an initial feel for the *relative* strengths and weaknesses of each instrument/algorithm.

- To provide useful input into the validation working group on the best levels of agreement we might expect to see between instruments in the 'ideal world'.
- To start to "speak each other's language".
- To identify and solve science and data interchange issues now rather than after launch.
- To look forward to future scientific studies involving data from multiple instruments.

It was also felt that if there was to be a push towards having 'definitive products' from Aura, the results of such an exercise would be *vital* for its development.

3.1.2 Schedule and constraints

Given the workload the teams are working under individually, it is important to avoid this exercise taking up too much extra effort. The approach of the group was to choose a timescale over which we would like to perform the work, and scope the effort accordingly.

The appropriate amount of time was discussed, and the following schedule was identified.

- Start when all instrument teams have data output from their algorithms to be intercompared. Teams are aiming to have all reached this point by 1 May 2001
- From this point spend 5–8 months on the intercomparison work. Thus ending sometime in Fall/Winter 2001/2002, aiming for 1 November 2001.
- 2–3 months would then be spent on preparing and compiling a final report for the exercise. Aim to complete this 1 March 2002.

3.2 The input dataset

All the various species measured and/or needed by the Aura instruments were divided into three groups:

Those species measured by more than one instrument Temperature, O_3, H_2O, CH_4, CO, N_2O, HNO_3, NO_2, geopotential height, aerosol and cloud.

Those species measured or needed by more than on instrument (e.g. as a forward model contaminant) HCN(?), N_2O_5, $ClONO_2$, CF_2Cl_2, $CHCl_3$, BrO, SO_2, CCl_4, $CClF_3$, C_2H_6.

Those measured or needed by one instrument alone OH, HO_2, NO, HCl, HOCl, ClO, CH_3CN, CF_4, OCS, SF_6, NH_3, HCOOH.

While the simulation we envisage will not address all these species, it is important to ensure that if more than one instrument uses a given species that the same profiles are used by each. In the case of aerosol and clouds, it was decided that this first intercomparison exercise will consider aerosol and cloud free cases. Later studies will probably include these.

The input dataset will be a days worth of data, probably late Northern or Southern hemisphere winter / early spring. The field will be 3D from pole to pole, with a vertical extent from the surface to $\sim$80 km. In order to appropriately capture diurnal chemistry issues, the input data shall contain multiple fields within the day, separated by timesteps of order 20 minutes to one hour. The horizontal resolution shall be around 2 degrees square. The vertical resolution will be a function of altitude, with a highest resolution (sub 1 km) around the tropopause, with slightly poorer resolution ($\sim$2 km) in the lower troposphere and mid to upper stratosphere, and fairly coarse resolution ($\sim$4–6 km) in the mesosphere.

Following this point in the discussion, Doug Kinnison (NCAR) gave a talk on the MOZART model. Doug agreed to provide input data commensurate with our requirements above on an appropriate timescale. For those few species for which the model cannot produce data, the group will identify other data sources as the need arises.

3.3 Implementation of the algorithms

The approach chosen is a series of '1D' tests. In these tests, we pretend that each instrument views the same region of the atmosphere simultaneously, and that this region is horizontally homogeneous. In this manner we avoid additional uncertainty due to sampling differences. Later exercises will probably involve more 'complete' simulations of the instruments' observing geometry and spacecraft orbit using the same input data. A number (50–100) of profiles will be chosen from the model data, along some nominal orbit track, and will form the basis of the simulations.

Given these profiles, the instrument teams will generate model radiance fields, using their 'best' forward model algorithms, as opposed to the 'operational' forward model that will be used in the retrieval algorithms. Appropriate levels of noise will be added to these radiances before retrievals are attempted. If there is time during the intercomparison exercise, studies will be made of 'noise free' cases, but these are lower priority.

3.4 The intercomparison of results

Given the timescale chosen, and the amount of other work members of the group are involved in the exercise will be limited to the comparison of Temperature, ozone and water vapour, in the upper troposphere to mid stratosphere region. Of the species listed in section 3.2, therefore, the only ones of important consideration will be these three species, plus any others needed for their retrieval (e.g. N_2O?).

The intercomparison of the three species will be undertaken by identified groups of people for each species, with help from appropriate members of each science team. The comparison of temperature will be led by Helen Worden from TES (JPL). Water vapor intercomparison will be led by Hugh Pumphrey from MLS (Edinburgh University) for the stratosphere and Tony Clough from TES (AER in Boston) for the troposphere. The lead role in the ozone intercomparison will be identified soon (hopefully someone from the HIRDLS team).

The main metrics for the intercomparison will clearly be differences from the truth, and the estimated precisions. Other metrics such as averaging kernel widths, error reduction etc. will be considered in more detail during the exercise. The intercomparison will be performed on the surfaces common to all instruments, with linear interpolation in product (e.g. vmr, not log vmr) in log pressure space used as appropriate. Note that HIRDLS and TES retrieve log vmr rather than vmr, so there will be a slight conceptual inconsistency here. One exception to this may be tropospheric water vapor, where other interpolations may better reflect the behavior of the profile.

3.5 Practicalities

The input model data will arrive in NetCDF format. These will be interpolated to the locations of the profiles to be compared. They will then be output in HDF-EOS swath format, similar to the one used by HIRDLS.

The instrument teams will produce data for intercomparison in their respective Level 2 file formats. These all use HDF-EOS swath and conform to the guidelines shortly to be laid down by the DSWG.

A web/ftp site will be set up to coordinate some of these activities. Possible hosts include the Aura project office, or JPL. NCAR may have advantages due to the less stringent requirements on vetting of web pages.

3.6 Final results of exercise

A brief report will be produced from this exercise. This will be composed of a 3–4 page section for each species compared, mostly consisting of figures and tables. Nathaniel Livesey from MLS (group chair) will coordinate the report and write introductory and closing material. The document will also contain an appendix describing the 'lessons learned' from the exercise as regards data interchange and practical issues.

There will also be a presentation to the wider Aura science team at the meeting following the completion of the exercise. This will probably be the spring 2002 meeting in the Netherlands.

4 General discussion on last day

The main topic of general discussion was the group's response to the suggestion made at the spring 2000 Aura science team meeting, that there should be 'definitive products' from Aura.

There are a multiplicity of ways in which to obtain such products, and it was generally agreed that the best of these would probably involve data assimilation. This method is the only one which yields a product that is better than the sum of its parts. Other methods such as simple weighted averages are fraught with problems. For example, with TES having a duty cycle, the information content in this product would cycle from day to day.

It was generally agreed that it is probably too late to be considering producing such a product at launch, unless an organization such as the DAO were to take a lead in its production. The feeling of the group was that unless such a combined product is better than the sum of its parts, we should encourage users to choose data from a particular instrument most suited to their study.

5 Plans for later meetings

A subset of the group will probably meet at the fall 2000 'CHAPS' meeting. The Spring 2001 science team meeting will be our next main meeting. At this point, we will be about to begin the intercomparison exercise in earnest. The exercise itself will probably involve one or two small meetings between the section authors.